



1 3. Cancelled.

E 1 9 11
2 4. (Amended) The method of claim 6, upon determining that the sum is greater
3 than the long-term averaged energy and before determining the peak-to-mean likelihood
4 ratio, the method further comprises:
5 determining whether a difference between the long-term averaged energy and the
6 short-term averaged energy is less than a predetermined threshold;
7 determining that the current audio frame represents voice if the difference is greater
8 than the predetermined threshold; and
9 continuing by determining the peak-to-mean likelihood ratio if the difference is less
than the predetermined threshold.

1 5. (Amended) The method of claim 6, wherein the determining of the short-term
2 averaged energy comprises:
3 determining an energy, in decibels, of the current audio frame;
4 determining a short-term averaged energy for a prior audio frame; and
5 conducting a weighted average of the energy of the current audio frame and the short-
6 term averaged energy for the prior audio frame.

E 2 11 12
3 6. (Three Times Amended) A method for enhancing voice activity detection
4 comprising:
5 determining a short-term averaged energy for a current audio frame;
6 determining a long-term averaged energy for the current audio frame;
7 determining whether a sum of the short-term averaged energy and a factor is greater
8 than the long-term averaged energy;

7 determining that the current audio frame represents silence if the sum is less than the
8 long-term averaged energy, without necessitating a determination of the peak-to-mean
9 likelihood ratio;

10 determining a peak-to-mean likelihood ratio, the determining a peak-to-mean
11 likelihood ratio comprises

12 calculating an averaged peak-to-mean ratio for the current audio frame,

13 determining a maximum averaged peak-to-mean ratio,

14 determining a minimum averaged peak-to-mean ratio,

15 determining a difference between the maximum averaged peak-to-mean ratio
16 and the averaged peak-to-mean ratio for the current audio frame,

17 determining a difference between the maximum averaged peak-to-mean ratio
18 and the minimum averaged peak-to-mean ratio, and

19 conducting a ratio, a denominator of the ratio being the difference between the
20 maximum averaged peak-to-mean ratio and the minimum averaged peak-to-mean
21 ratio, the numerator being the difference between the maximum averaged peak-to-
22 mean ratio and the averaged peak-to-mean ratio; and

23 comparing the peak-to-mean likelihood ratio to a selected threshold to determine
24 whether the current audio frame represents a voice signal.

1 7. Cancelled.

1 8. Cancelled.

E3 ~~9. (Amended) The communication module of claim 12, wherein the voice
2 activity detector, when executed, controls the processing unit to determine whether a
3 difference between the long-term averaged energy and the short-term averaged energy is less~~

E3
4 than a predetermined threshold, and to signal that the current audio frame represents voice if
5 the difference is greater than the predetermined threshold.

1 10. Cancelled.

E4
11. (Amended) The communication module of claim 9, wherein the voice activity
2 detector, when executed, controls the processing unit to determine a peak-to-mean ratio by (i)
3 sampling an analog signal a predetermined number of times to produce a plurality of sampled
4 signals each having a sampled value, (ii) determining a maximum value of the plurality of
5 sampled signals, and (iii) conducting a ratio between an absolute value of the maximum
6 value and a summation of the sampled values for the plurality of sampled signals.

E5
12. (Three Times Amended) A communication module comprising:
a substrate;
a processing unit placed on the substrate; and
a memory coupled to the processing unit, the memory to contain a voice activity
detector which, when executed, controls the processing unit to
determine whether a sum of a short-term averaged energy and a
predetermined factor is greater than a long-term averaged energy, and to signal that a
current audio frame represents silence if the sum is less than the long-term averaged
energy, and
if the current audio frame is not determined to be silence using the short-term
averaged energy and the long-term averaged energy, determine a peak-to-mean
likelihood ratio for the current audio frame by (i) monitoring a maximum averaged
peak-to-mean ratio and a minimum averaged peak-to-mean ratio, (ii) determining a
first result being a difference between the maximum averaged peak-to-mean ratio and
the averaged peak-to-mean ratio for the current audio frame, (iii) determining a
second result being a difference between the maximum averaged peak-to-mean ratio

E5
cond

18 and the minimum averaged peak-to-mean ratio, and (iv) conducting a ratio between
19 the first result as a numerator and the second result as a denominator, and comparing
20 the peak-to-mean likelihood ratio to a selected threshold to determine whether the
current audio frame represents a voice signal.

1 13. Cancelled.

1 14. Cancelled.

7/6 52/3

15. (Amended) A machine readable medium having embodied thereon a
computer program for processing by a machine, the computer program comprising:
a first routine for determining a normalized peak-to-mean likelihood ratio including
4 (i) a denominator having a value substantially equal to a difference between a maximum
5 averaged peak-to-mean ratio and a minimum averaged peak-to-mean ratio and (ii) a
6 numerator having a value substantially equal to a difference between the maximum averaged
7 peak-to-mean ratio and the averaged peak-to-mean ratio;
8 a second routine for comparing the peak-to-mean likelihood ratio to a selected
9 threshold to determine whether a current audio frame being transmitted represents a voice
10 signal;
11 a third routine for determining a short-term averaged energy for successive audio
12 frames including the current audio frame, the third routine being executed before the first and
13 second routines;
14 a fourth routine for determining a long-term averaged energy for the current audio
15 frame, the fourth routine being executed before the first and second routines;

16 a fifth routine for determining whether a sum of the short-term averaged energy and a
17 predetermined factor is greater than the long-term averaged energy, the fifth routine being
18 executed before the first and second routines; and
19 a sixth routine for determining whether a difference between the long-term averaged
20 energy and the short-term averaged energy is less than a predetermined threshold, the sixth
21 routine being executed after determining that the sum is greater than the long-term averaged
22 energy and before execution of the first and second routines.

1 16. The machine readable medium of claim 15, wherein the fifth routine
2 determining that the current audio frame represents silence if the sum is less than the long-
3 term averaged energy.

1 17. The machine readable medium of claim 15, wherein the sixth routine
2 determining that the current audio frame represents voice if the difference is greater than the
3 predetermined threshold.

1 18. Cancelled.

1 19. Cancelled.

1 20. Cancelled.

1 21. Cancelled.

EX 22. (Amended) A method for enhancing voice activity detection comprising:
2 determining a short-term averaged energy for a current audio frame;
3 determining a long-term averaged energy for the current audio frame;
4 determining whether a sum of the short-term averaged energy and a factor is greater
5 than the long-term averaged energy;

SA 6
E 7
C 1
7
determining that the current audio frame represents silence if the sum is less than the
long-term averaged energy, without necessitating a determination of the peak-to-mean
likelihood ratio;
determining a peak-to-mean likelihood ratio including (i) a denominator having a
value substantially equal to a difference between a maximum averaged peak-to-mean ratio
and a minimum averaged peak-to-mean ratio and (ii) a numerator having a value
substantially equal to a difference between the maximum averaged peak-to-mean ratio and
the averaged peak-to-mean ratio; and
comparing the peak-to-mean likelihood ratio to a selected threshold to determine
whether a current audio frame represents a voice signal.

23. The method of claim 22, upon determining that the sum is greater than the
long-term averaged energy and before determining the peak-to-mean likelihood ratio, the
method further comprises:
determining whether a difference between the long-term averaged energy and the
short-term averaged energy is less than a predetermined threshold;
determining that the current audio frame represents voice if the difference is greater
than the predetermined threshold; and
continuing by determining the peak-to-mean likelihood ratio if the difference is less
than the predetermined threshold.

EX 1
2
3
4
24. (Amended) The method of claim 22, wherein the determining of the short-
term averaged energy comprises:
determining an energy, in decibels, of the current audio frame;
determining a short-term averaged energy for a prior audio frame; and

E8
con
SIX
conducting a weighted average of the energy of the current audio frame and the short-
6 term averaged energy for the prior audio frame.

E9
SIX
25. (New) The method of claim 6, wherein the short-term averaged energy is an
4 accumulation of signal energy associated with successive audio frames including the current
3 audio frame.

1 26. (New) The method of claim 25, wherein the successive audio frames are
2 pulse code modulation (PCM) audio frames.

1 27. (New) The method of claim 25, wherein the long-term averaged energy is
2 based on the accumulation of the signal energy and a background noise level.

1 28. (New) The method of claim 6, wherein the short-term averaged energy is
2 based on a current frame entry and a prior short-term averaged energy value.

1 29. (New) The method of claim 6, wherein the factor is at least two decibels.

1 30. (New) The communication module of claim 12, wherein the short-term
2 averaged energy determined by the voice activity detector is an accumulation of signal
3 energy associated with the successive audio frames being pulse code modulation (PCM)
4 audio frames.

1 31. (New) The communication module of claim 30, wherein the long-term
2 averaged energy determined by the voice activity detector is based on the accumulation of
3 the signal energy and a background noise level.

1 32. (New) The communication module of claim 12, wherein the predetermined
2 factor is at least two decibels.

1 33. (New) The software readable medium of claim 15, wherein the short-term
2 averaged energy determined by the third routine is an accumulation of signal energy
3 associated with the successive audio frames.

509
74
E9
Lone
1 34. (New) The software readable medium of claim 33, wherein the long-term
2 averaged energy determined by the fourth routine is based on the accumulation of the signal
3 energy and a background noise level.

1 35. (New) The method of claim 22, wherein the short-term averaged energy is an
2 accumulation of signal energy associated with successive audio frames including the current
3 audio frame.

1 36. (New) The method of claim 22, wherein the short-term averaged energy is
2 based on the current audio frame and a prior short-term averaged energy value.

1 37. (New) The method of claim 22, wherein the factor is at least two decibels.